Lecture 9.

Let F be a finite field, $|F| = q$. A code C is a subset of the vector space F^n. We refer to it as a q-ary code. If $q=2$ i.e. $F = \mathbb{Z}(2)$ then the code is binary, if $|F| = 3$ i.e. $F = \mathbb{Z}(3)$ then the code is ternary etc.

Given two vectors $v = (\alpha_1, \dots, \alpha_n)$ and $w = (\beta_1, \dots, \beta_n)$ the Hamming distance is defined as

$$d(v, w) = |\{ i \mid \alpha_i \neq \beta_i \}|.$$

The Hamming distance makes F^n a metric space:

1) $d(v,w) \geq 0$, $d(v,w) = 0 \iff v = w$.

2) $d(v,w) = d(w,v)$.

3) (triangle inequality) $d(v,w) \leq d(v,u) + d(u,w)$.

A *Hamming ~~sphere~~ ball* of radius r with the center at $v \in F^n$ is defined as

$$B(v,r) = \{ w \in F^n \mid d(v,w) \leq r \}.$$

Let $0 \leq i \leq n$. It is easy to see that

$$|\{ w \mid d(v,w) = i \}| = \binom{n}{i}(q-1)^i$$

Hence $|B(v,r)| = \sum_{i=0}^{r} \binom{n}{i}(q-1)^i$.

Def. The *Hamming weight* of a code C is defined as

$$d = \min\{ d(v,w) \mid v, w \in C, v \neq w \}.$$

Lemma. Let $r = \left[\frac{d-1}{2} \right]$. All balls $B(v,r)$, $v \in C$, are disjoint.

Proof. Suppose that $v \neq w$; $v, w \in C$ and $B(v, r) \cap B(w, r) \ni u$.

It means that $d(v, u) \leq r$, $d(u, w) \leq r$.

By the Triangle Inequality

$$d(v, w) \leq d(v, u) + d(u, w) \leq 2r =$$

$$2\left[\frac{d-1}{2}\right] \leq 2 \cdot \frac{d-1}{2} = d-1,$$

which contradicts minimality of d. I

COROLLARY. Assume that we make $\leq \left[\frac{d-1}{2}\right]$ mistakes. Then C uniquely determines the message:

if we got the message u then there is a unique codeword $w \in C$ such that

$$d(u, w) \leq \left[\frac{d-1}{2}\right].$$

If there were two such codewords $v \neq w$ ~~Such~~ then $u \in B(v, [\frac{d-1}{2}]) \cap B(w, [\frac{d-1}{2}])$, a contradiction.

In other words: having received u we choose $w \in C$ that is closest to u.

Ex. Is there $C \subset \mathbb{Z}(2)^7$, $|C| = 17$, $d = 3$?

$[\frac{d-1}{2}] = 1$.

$|B(v, 1)| = 1 + \binom{7}{1}(2-1) = 8$

$17 \cdot 8 = 136$. But $|\mathbb{Z}(2)^7| = 2^7 = 128$.

Hence there is no such code.

Def. C is a *linear code* if C is a subspace of a vector space F^n.

From now on we talk only about

linear codes.

Let $\dim_F C = k$. Then we refer to C as a $[n,k]$-code.

<u>Def.</u> A $k \times n$ matrix

$$G = \left(\underbrace{\begin{matrix} \rule{4em}{0.4pt} \\ \rule{4em}{0.4pt} \\ \cdots \\ \rule{4em}{0.4pt} \end{matrix}}_{n} \right) \Big\} k,$$

whose rows form a basis of C is called a <u>generator matrix</u> of C.

<u>Remark</u>- A generator matrix is not unique.

<u>Def.</u> If C is a linear code then

$$C^{\perp} = \{ v \in F^n \mid v \cdot C = (0) \}$$

is the <u>dual code</u> of C.

If $v=(\alpha_1,\dots,\alpha_n)$, $w=(\beta_1,\dots,\beta_n)$ then

$$v\cdot w=\alpha_1\beta_1+\cdots+\alpha_n\beta_n.$$

$$C^\perp=\{v\in F^n \mid v\cdot w=0 \;\forall w\in C\}.$$

<u>Def.</u> A generator matrix H of $C^\perp$ is called a <u>parity check matrix</u> of C,

$$C=\{v\in F^n \mid Hv^T=0\}.$$

<u>Def.</u> The <u>Hamming weight</u> of $v\in F^n$ is

$wt(v)=\{i \mid v_i\neq 0\}=d(v,0)$,

Hamming weight of C

$$d(C)=\min\{wt(v) \mid 0\neq v\in C\}$$

We refer to the linear code C as a $[n,k,d]$-code.

<u>Example.</u> For $C=F^n$ C is an $[n,n,1]$-code.

For $v, w \in \mathbb{Z}(2)^n$ let

$$v * w = (v_1 w_1, \dots, v_n w_n),$$

here $v = (v_1, \dots, v_n)$, $w = (w_1, \dots, w_n)$.

Then $$wt(v+w) = wt(v) + wt(w) + 2\, wt(v*w)$$

Example. Let C be a binary linear code,

$C = \{ v \in \mathbb{Z}(2)^n \mid wt(v) \text{ is even} \}$.

Exercise : prove that $\dim C = n-1$, $d(C) = 2$.

Find a generator and a parity check matrices for C.

Proposition. The Hamming weight of a linear code C with a parity check matrix H is equal to the largest integer d

such that every $d-1$ columns of H are linearly indgendent.

<u>Proof</u>. If every $d-1$ columns of H are linearly indgendent, then C does not contain vectors of weight k, $k \le d-1$. Indeed, let $H = [\bar{h}_1 \cdots \bar{h}_n]$, $\bar{h}_j$ are columns of H. Let $v \in C$, $v = (0 \cdots \alpha_{i_1} \cdots \alpha_{i_k} \cdots 0)$, where $\alpha_{i_1}, \dots, \alpha_{i_k} \neq 0$, $k \le d-1$.

Then
$$H v^T = [\bar{h}_1 \cdots \bar{h}_n] \begin{pmatrix} 0 \\ \vdots \\ \alpha_{i_1} \\ \vdots \\ \alpha_{i_k} \\ 0 \\ \vdots \\ 0 \end{pmatrix} =$$

$$\alpha_{i_1} \bar{h}_{i_1} + \cdots + \alpha_{i_k} \bar{h}_{i_k} = 0,$$

which means that $\bar{h}_{i_1}, \dots, \bar{h}_{i_k}$ are linearly dependent.

Since d is maximal there are d columns $\bar{h}_{j_1}, \dots, \bar{h}_{j_d}$ that are linearly dependent. Let

$\alpha_{j_1} \bar{h}_{j_1} + \dots + \alpha_{j_d} \bar{h}_{j_d}$; $\alpha_{j_1}, \dots, \alpha_{j_d} \neq 0$.

Then $(0 \cdots \alpha_{j_1} \cdots \alpha_{j_d} \cdots 0) \in C$, $d =$ Hamming weight of C. $\rfloor$

Example. $H = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$.

Any 2 columns are linearly independent. But the 1s, 3d, 5th columns are linearly dependent. Hence H is a parity check matrix of a $[5, 2, 3]$-code.

Theorem (The Singleton Bound). If C is a $[n, k, d]$-code then

$$d \leq n-k+1.$$

<u>Proof</u>. Let H be a $(n-k) \times n$ parity check matrix of C. The rank of H is $n-k$. Hence every $n-k+1$ columns are linearly dependent. ∇

<u>Def</u>. An $[n,k,d]$-code with $d = n-k+1$ is called a <u>maximum distance separable code</u> or an MDS-code.

<u>Example</u>. $C = \{ v = (\alpha_1, \dots, \alpha_n) \,|\, wt\ v \text{ is even} \}$. Then $k = n-1$, $d = 2 = n-k+1$. So it is an MDS-code.

<u>Def</u>. A generator matrix of the form $[I_k \,|\, B]$ is called a <u>standard generator matrix</u>.

The permutation group $\underline{P}_n$ acts on F^n by permuting components

$$(ij)(\alpha_1 \cdot\cdot \alpha_i \cdot\cdot \alpha_j \cdot\cdot \alpha_n) = (\alpha_1 \cdots \alpha_j \cdots \alpha_i \cdots \alpha_n).$$

the action by $g \in \underline{P}_n$ is a linear transformation of F^n.

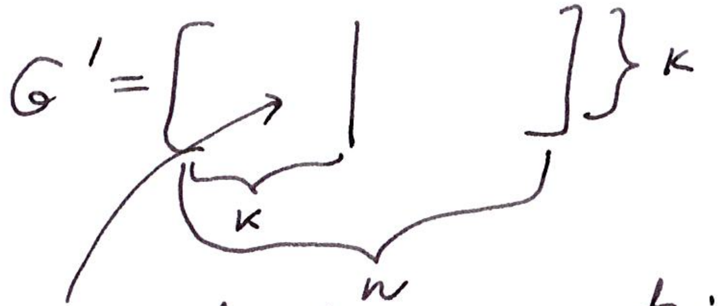
Def. Two codes C, C' are called *permutation equivalent* if there exists $g \in \underline{P}_n : gC = C'$.

Proposition. Every linear code is permutation equivalent to a code with a standard generator matrix.

Proof. Let $G = \left.\underbrace{\begin{bmatrix} \qquad & \qquad \\ \qquad & \qquad \end{bmatrix}}_{n}\right\} k$ be a generator matrix of a code C. The rank of G is k.

Applying a permutation of columns we change G to a matrix G' having the first k columns linearly independent.

The code $C \rightsquigarrow$ code C'



nonsingular matrix.

<u>A fact from Linear Algebra</u>: every nonsingular $k \times k$ matrix can be reduced to I_k by elementary transformations of rows. Elementary transformations:

(1) multiplication of a row by $\alpha \neq 0$,

(2) adding one row $\times \alpha$ to another row.

these elementary transformations of rows amount to a change of basis in C'. They do not change C'! Hence C' has a generator matrix of the form $(I_k | B)$ ⊥

<u>Def.</u> A set of k positions $i_1, \dots, i_k$ of an $[n,k]$-code is called an <u>information set</u> if the i_1-th, i_2-th, …, i_k-th columns of G are linearly independent.

<u>Exercise.</u> Prove that this definition does not depend on a choice of a basis in C.

<u>Remark.</u> C admits a standard generator matrix if and only if the first k positions $1, 2, \dots, k$ is an information set.